

0590
0424

H3

OIKE

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/10/040,416

TIME: 14:32:22

Input Set : A:\217969US0XCIP.txt

Output Set: N:\CRF3\04292002\J040416.raw

p6

3 <110> APPLICANT: OOKURA, TETUYA
 4 KASUMI, TAKAFUMI
 5 ASABA, EIJI
 7 <120> TITLE OF INVENTION: ERYTHROSE REDUCTASE, ITS cDNA AND CELL WHICH THE cDNA
 EXPRESS

9 <130> FILE REFERENCE: 217969US0XCIP
 11 <140> CURRENT APPLICATION NUMBER: 10/040,416
 12 <141> CURRENT FILING DATE: 2002-01-09
 14 <150> PRIOR APPLICATION NUMBER: US 09/800,487
 15 <151> PRIOR FILING DATE: 2001-03-08
 17 <150> PRIOR APPLICATION NUMBER: JP 2001-001294
 18 <151> PRIOR FILING DATE: 2001-01-09
 20 <160> NUMBER OF SEQ ID NOS: 10
 22 <170> SOFTWARE: PatentIn version 3.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 1119
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Trichosporonoides megachiliensis
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (1)..(993)
 32 <223> OTHER INFORMATION:

ENTERED

35 <400> SEQUENCE: 1
 36 atg tct tac aaa cag tac atc ccc ctg aac gac ggt aac aaa atc cct 48
 37 Met Ser Tyr Lys Gln Tyr Ile Pro Leu Asn Asp Gly Asn Lys Ile Pro
 38 1 5 10 15
 40 gcc ctt gga ttt ggt act tgg caa gct gaa cct ggt caa gtg ggt gca 96
 41 Ala Leu Gly Phe Gly Thr Trp Gln Ala Glu Pro Gly Gln Val Gly Ala
 42 20 25 30
 44 agt gtc aag aac gct gtc aag gct ggg tac cgt cat ttg gat ttg gcc 144
 45 Ser Val Lys Asn Ala Val Lys Ala Gly Tyr Arg His Leu Asp Leu Ala
 46 35 40 45
 48 aaa gtg tac caa aac caa tcg gaa att gga gta gca ctt cag gaa ctg 192
 49 Lys Val Tyr Gln Asn Gln Ser Glu Ile Gly Val Ala Leu Gln Glu Leu
 50 50 55 60
 52 ttt gat caa ggt att gtt aaa cgg gaa gat ttg ttt att acg tcc aaa 240
 53 Phe Asp Gln Gly Ile Val Lys Arg Glu Asp Leu Phe Ile Thr Ser Lys
 54 65 70 75 80
 56 gta tgg aat aac cgt cat gct cct gaa cat gtt gag cct gca ttg gac 288
 57 Val Trp Asn Asn Arg His Ala Pro Glu His Val Glu Pro Ala Leu Asp
 58 85 90 95
 60 gaa aca ttg aaa gaa ctt gga ttg tcc tac ttg gat ttg tac ttg att 336
 61 Glu Thr Leu Lys Glu Leu Gly Leu Ser Tyr Leu Asp Leu Tyr Leu Ile
 62 100 105 110

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/10/040,416

TIME: 14:32:22

Input Set : A:\217969US0XCIP.txt

Output Set: N:\CRF3\04292002\J040416.raw

```

64 cat tgg ccc gtt gcg ttc aag ttt act acg cct caa gaa cta ttc cct      384
65 His Trp Pro Val Ala Phe Lys Phe Thr Thr Pro Gln Glu Leu Phe Pro
66      115      120      125
68 act gag ccg gat aac aag gaa ttg gcc gcg att gat gat tca atc aag      432
69 Thr Glu Pro Asp Asn Lys Glu Leu Ala Ala Ile Asp Asp Ser Ile Lys
70      130      135      140
72 ttg gta gac act tgg aag gca gtt gta gca ctc aaa acg ggt aag      480
73 Leu Val Asp Thr Trp Lys Ala Val Val Ala Leu Lys Lys Thr Gly Lys
74 145      150      155      160
76 acc aaa tcc gtt ggt gtg tcg aac ttc act acg gat ttg gta gac ttg      528
77 Thr Lys Ser Val Gly Val Ser Asn Phe Thr Thr Asp Leu Val Asp Leu
78      165      170      175
80 gtt gaa aaa gcg tcg ggg gaa cga ccg gcg gtc aat cag att gaa gca      576
81 Val Glu Lys Ala Ser Gly Glu Arg Pro Ala Val Asn Gln Ile Glu Ala
82      180      185      190
84 cac cca ttg tta caa cag gat gaa ttg gtt gct cat cac aag agt aaa      624
85 His Pro Leu Leu Gln Gln Asp Glu Leu Val Ala His His Lys Ser Lys
86      195      200      205
88 aac att gtg att act gcg tac agt cct ttg gga aac aat gtg agt ggg      672
89 Asn Ile Val Ile Thr Ala Tyr Ser Pro Leu Gly Asn Asn Val Ser Gly
90      210      215      220
92 aaa cca cct ctg act caa aac cct ggg att gaa gca act gcg aaa cgg      720
93 Lys Pro Pro Leu Thr Gln Asn Pro Gly Ile Glu Ala Thr Ala Lys Arg
94 225      230      235      240
96 tta aat cat act cct gct gcg gtc ttg ctt gca tgg ggg att caa cgt      768
97 Leu Asn His Thr Pro Ala Ala Val Leu Leu Ala Trp Gly Ile Gln Arg
98      245      250      255
100 gga tac agt gta ttg gtc aag agt gtt aca cct tct cga att gag agc      816
101 Gly Tyr Ser Val Leu Val Lys Ser Val Thr Pro Ser Arg Ile Glu Ser
102      260      265      270
104 aat tat gat cag att acc ctt tct cct gaa gaa ttc cag aag gtt acg      864
105 Asn Tyr Asp Gln Ile Thr Leu Ser Pro Glu Glu Phe Gln Lys Val Thr
106      275      280      285
108 gat ttg atc aag gaa tat ggc gaa agt cgc aac aat att ccg ttg aat      912
109 Asp Leu Ile Lys Glu Tyr Gly Glu Ser Arg Asn Asn Ile Pro Leu Asn
110      290      295      300
112 tat aaa cct tca tgg ccc atc agt gtg ttt ggt aca tcg gat gaa gct      960
113 Tyr Lys Pro Ser Trp Pro Ile Ser Val Phe Gly Thr Ser Asp Glu Ala
114 305      310      315      320
116 aag gct act cat aag att aac acc aac ctt tga gttcagtttg ggaactattt      1013
117 Lys Ala Thr His Lys Ile Asn Thr Asn Leu
118      325      330
120 aaagctgctt gctggtcaca ttattgtcag tacctaccat gaagaattca atattatttt      1073
122 acattgtcaa ccattacatg gatccaaaaa aaaaaaaaaa aaaaaa      1119
125 <210> SEQ ID NO: 2
126 <211> LENGTH: 330
127 <212> TYPE: PRT
128 <213> ORGANISM: Trichosporonoides megachiliensis
130 <400> SEQUENCE: 2

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/040,416

DATE: 04/29/2002

TIME: 14:32:22

Input Set : A:\217969US0XCIP.txt

Output Set: N:\CRF3\04292002\J040416.raw

```

132 Met Ser Tyr Lys Gln Tyr Ile Pro Leu Asn Asp Gly Asn Lys Ile Pro
133 1          5          10          15
136 Ala Leu Gly Phe Gly Thr Trp Gln Ala Glu Pro Gly Gln Val Gly Ala
137          20          25          30
140 Ser Val Lys Asn Ala Val Lys Ala Gly Tyr Arg His Leu Asp Leu Ala
141          35          40          45
144 Lys Val Tyr Gln Asn Gln Ser Glu Ile Gly Val Ala Leu Gln Glu Leu
145          50          55          60
148 Phe Asp Gln Gly Ile Val Lys Arg Glu Asp Leu Phe Ile Thr Ser Lys
149 65          70          75          80
152 Val Trp Asn Asn Arg His Ala Pro Glu His Val Glu Pro Ala Leu Asp
153          85          90          95
156 Glu Thr Leu Lys Glu Leu Gly Leu Ser Tyr Leu Asp Leu Tyr Leu Ile
157          100          105          110
160 His Trp Pro Val Ala Phe Lys Phe Thr Thr Pro Gln Glu Leu Phe Pro
161          115          120          125
164 Thr Glu Pro Asp Asn Lys Glu Leu Ala Ala Ile Asp Asp Ser Ile Lys
165          130          135          140
168 Leu Val Asp Thr Trp Lys Ala Val Val Ala Leu Lys Lys Thr Gly Lys
169 145          150          155          160
172 Thr Lys Ser Val Gly Val Ser Asn Phe Thr Thr Asp Leu Val Asp Leu
173          165          170          175
176 Val Glu Lys Ala Ser Gly Glu Arg Pro Ala Val Asn Gln Ile Glu Ala
177          180          185          190
180 His Pro Leu Leu Gln Gln Asp Glu Leu Val Ala His His Lys Ser Lys
181          195          200          205
184 Asn Ile Val Ile Thr Ala Tyr Ser Pro Leu Gly Asn Asn Val Ser Gly
185          210          215          220
188 Lys Pro Pro Leu Thr Gln Asn Pro Gly Ile Glu Ala Thr Ala Lys Arg
189 225          230          235          240
192 Leu Asn His Thr Pro Ala Ala Val Leu Leu Ala Trp Gly Ile Gln Arg
193          245          250          255
196 Gly Tyr Ser Val Leu Val Lys Ser Val Thr Pro Ser Arg Ile Glu Ser
197          260          265          270
200 Asn Tyr Asp Gln Ile Thr Leu Ser Pro Glu Glu Phe Gln Lys Val Thr
201          275          280          285
204 Asp Leu Ile Lys Glu Tyr Gly Glu Ser Arg Asn Asn Ile Pro Leu Asn
205          290          295          300
208 Tyr Lys Pro Ser Trp Pro Ile Ser Val Phe Gly Thr Ser Asp Glu Ala
209 305          310          315          320
212 Lys Ala Thr His Lys Ile Asn Thr Asn Leu
213          325          330
216 <210> SEQ ID NO: 3
217 <211> LENGTH: 1077
218 <212> TYPE: DNA
219 <213> ORGANISM: Trichosporonoides megachiliensis
221 <220> FEATURE:
222 <221> NAME/KEY: CDS
223 <222> LOCATION: (1)..(987)

```

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/10/040,416

TIME: 14:32:22

Input Set : A:\217969US0XCIP.txt

Output Set: N:\CRF3\04292002\J040416.raw

224 <223> OTHER INFORMATION:

227 <400> SEQUENCE: 3

228	atg	tcc	tac	aac	aag	aac	atc	cct	ctc	aac	gac	ggt	aac	tcc	att	cct	48
229	Met	Ser	Tyr	Asn	Lys	Asn	Ile	Pro	Leu	Asn	Asp	Gly	Asn	Ser	Ile	Pro	
230	1			5						10					15		
232	gct	ctt	ggg	tac	ggt	acc	tgg	caa	gca	gaa	cct	ggt	cag	gtc	ggt	gaa	96
233	Ala	Leu	Gly	Tyr	Gly	Thr	Trp	Gln	Ala	Glu	Pro	Gly	Gln	Val	Gly	Glu	
234				20					25					30			
236	ggt	gtc	aag	ctc	gct	gtt	aag	gct	ggc	tac	cgt	cat	ttg	gac	ttg	gcc	144
237	Gly	Val	Lys	Leu	Ala	Val	Lys	Ala	Gly	Tyr	Arg	His	Leu	Asp	Leu	Ala	
238			35				40					45					
240	aaa	gta	tac	cag	aac	caa	acc	gag	att	ggc	caa	gct	ctc	aag	gaa	ctg	192
241	Lys	Val	Tyr	Gln	Asn	Gln	Thr	Glu	Ile	Gly	Gln	Ala	Leu	Lys	Glu	Leu	
242		50				55				60							
244	ttt	gat	gag	ggt	gtt	gtc	aag	cgt	gag	gac	ctt	ttc	atc	act	tcc	aag	240
245	Phe	Asp	Glu	Gly	Val	Val	Lys	Arg	Glu	Asp	Leu	Phe	Ile	Thr	Ser	Lys	
246	65				70				75						80		
248	ctt	tgg	aac	aac	cgc	cac	gct	cct	gag	cac	gtt	gag	cct	gcg	ctc	gac	288
249	Leu	Trp	Asn	Asn	Arg	His	Ala	Pro	Glu	His	Val	Glu	Pro	Ala	Leu	Asp	
250				85					90					95			
252	gag	act	ctt	aag	gag	ttg	ggt	cta	tcc	tat	ttg	gac	ctg	tac	ttg	att	336
253	Glu	Thr	Leu	Lys	Glu	Leu	Gly	Leu	Ser	Tyr	Leu	Asp	Leu	Tyr	Leu	Ile	
254			100					105					110				
256	cac	tgg	cct	gtt	gct	ttc	aag	ttc	act	act	ccc	gat	gaa	ctg	ctt	cct	384
257	His	Trp	Pro	Val	Ala	Phe	Lys	Phe	Thr	Thr	Pro	Asp	Glu	Leu	Leu	Pro	
258			115				120					125					
260	gct	gac	cct	acc	aac	aag	gac	ctt	gcc	tac	att	gac	gat	tcg	gtc	aaa	432
261	Ala	Asp	Pro	Thr	Asn	Lys	Asp	Leu	Ala	Tyr	Ile	Asp	Asp	Ser	Val	Lys	
262		130				135						140					
264	ttg	tcc	gac	acc	tgg	aag	gcg	gtc	gtc	gcc	ctg	aaa	aag	acg	ggt	aag	480
265	Leu	Ser	Asp	Thr	Trp	Lys	Ala	Val	Val	Ala	Leu	Lys	Lys	Thr	Gly	Lys	
266	145				150				155						160		
268	acc	aag	tcg	gtt	ggt	gtt	tcg	aac	ttc	agc	act	cgt	ctg	gtc	gac	ttg	528
269	Thr	Lys	Ser	Val	Gly	Val	Ser	Asn	Phe	Ser	Thr	Arg	Leu	Val	Asp	Leu	
270				165					170					175			
272	gtt	gag	gaa	gct	tcg	ggc	gaa	cgt	cct	gcg	gtt	aac	cag	atc	gaa	gct	576
273	Val	Glu	Glu	Ala	Ser	Gly	Glu	Arg	Pro	Ala	Val	Asn	Gln	Ile	Glu	Ala	
274			180					185					190				
276	cac	ccc	ttg	ttg	caa	caa	gac	gag	ttg	gtt	gct	cac	cat	aag	agc	aag	624
277	His	Pro	Leu	Leu	Gln	Gln	Asp	Glu	Leu	Val	Ala	His	His	Lys	Ser	Lys	
278			195				200					205					
280	aac	att	gtc	atc	act	gct	tac	agt	ccc	ttg	ggc	aac	aat	gtc	gct	ggt	672
281	Asn	Ile	Val	Ile	Thr	Ala	Tyr	Ser	Pro	Leu	Gly	Asn	Asn	Val	Ala	Gly	
282		210				215						220					
284	aaa	cca	cct	ctg	act	gag	aat	ccc	ggt	att	gtg	gat	gct	gct	aag	cgt	720
285	Lys	Pro	Pro	Leu	Thr	Glu	Asn	Pro	Gly	Ile	Val	Asp	Ala	Ala	Lys	Arg	
286	225				230				235					240			
288	ctg	aac	cat	act	cct	gct	gct	gtg	ctc	att	gct	tgg	ggt	att	caa	cgc	768
289	Leu	Asn	His	Thr	Pro	Ala	Ala	Val	Leu	Ile	Ala	Trp	Gly	Ile	Gln	Arg	

RAW SEQUENCE LISTING

DATE: 04/29/2002

PATENT APPLICATION: US/10/040,416

TIME: 14:32:22

Input Set : A:\217969US0XCIP.txt

Output Set: N:\CRF3\04292002\J040416.raw

```

290          245          250          255
292 ggg tac agc gtc ttg gtc aag tca gtt aca ccc tca cgg att aag agt      816
293 Gly Tyr Ser Val Leu Val Lys Ser Val Thr Pro Ser Arg Ile Lys Ser
294          260          265          270
296 aac ttt gaa cag atc act ctg tct gat gag gaa ttc caa cgg gtt acc      864
297 Asn Phe Glu Gln Ile Thr Leu Ser Asp Glu Glu Phe Gln Arg Val Thr
298          275          280          285
300 aac ctc atc aag gag tac ggt gag agc cgt aac aac gtt cct ttc aac      912
301 Asn Leu Ile Lys Glu Tyr Gly Glu Ser Arg Asn Asn Val Pro Phe Asn
302          290          295          300
304 tac aag cct tcg tgg tct att gac gtc ttt ggt acc cag tac gag gct      960
305 Tyr Lys Pro Ser Trp Ser Ile Asp Val Phe Gly Thr Gln Tyr Glu Ala
306 305          310          315          320
308 aag gct acc cac aag att aac gct taa tgtgctctta tcaaaaaagt      1007
309 Lys Ala Thr His Lys Ile Asn Ala
310          325
312 cgtttgaacc tgtaatgtgt gaatgttatc ctcattgttg catcatctca tcaaaaaaaaa      1067
314 aaaaaaaaaa      1077
317 <210> SEQ ID NO: 4
318 <211> LENGTH: 328
319 <212> TYPE: PRT
320 <213> ORGANISM: Trichosporonoides megachiliensis
322 <400> SEQUENCE: 4
324 Met Ser Tyr Asn Lys Asn Ile Pro Leu Asn Asp Gly Asn Ser Ile Pro
325 1          5          10          15
328 Ala Leu Gly Tyr Gly Thr Trp Gln Ala Glu Pro Gly Gln Val Gly Glu
329          20          25          30
332 Gly Val Lys Leu Ala Val Lys Ala Gly Tyr Arg His Leu Asp Leu Ala
333          35          40          45
336 Lys Val Tyr Gln Asn Gln Thr Glu Ile Gly Gln Ala Leu Lys Glu Leu
337          50          55          60
340 Phe Asp Glu Gly Val Val Lys Arg Glu Asp Leu Phe Ile Thr Ser Lys
341 65          70          75          80
344 Leu Trp Asn Asn Arg His Ala Pro Glu His Val Glu Pro Ala Leu Asp
345          85          90          95
348 Glu Thr Leu Lys Glu Leu Gly Leu Ser Tyr Leu Asp Leu Tyr Leu Ile
349          100          105          110
352 His Trp Pro Val Ala Phe Lys Phe Thr Thr Pro Asp Glu Leu Leu Pro
353          115          120          125
356 Ala Asp Pro Thr Asn Lys Asp Leu Ala Tyr Ile Asp Asp Ser Val Lys
357          130          135          140
360 Leu Ser Asp Thr Trp Lys Ala Val Val Ala Leu Lys Lys Thr Gly Lys
361 145          150          155          160
364 Thr Lys Ser Val Gly Val Ser Asn Phe Ser Thr Arg Leu Val Asp Leu
365          165          170          175
368 Val Glu Glu Ala Ser Gly Glu Arg Pro Ala Val Asn Gln Ile Glu Ala
369          180          185          190
372 His Pro Leu Leu Gln Gln Asp Glu Leu Val Ala His His Lys Ser Lys
373          195          200          205

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/040,416

DATE: 04/29/2002
TIME: 14:32:23

Input Set : A:\217969US0XCIP.txt
Output Set: N:\CRF3\04292002\J040416.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 9

Seq#:8; N Pos. 3,9,18